Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

The Reduction of Families of Bilinear Forms.

By H. E. Hawkes.

Introduction

The problem of determining the necessary and sufficient conditions that two families of non-singular bilinear forms are equivalent was first solved by Weierstrass in 1868. Since that time the subject has received considerable attention from various investigators, notably Frobenius, Kronecker, Weyr and Bromwich, but until recently a connected and clear account of this theory has been lacking. The work of Bromwich, and in greater measure Bôcher's Introduction to Higher Algebra, have served to supply the deficiency, so that the theory cannot now be called inaccessible.

The present paper gives a direct and practically self-contained derivation of the necessary and sufficient condition that two non-singular families of bilinear forms are equivalent by a method that is not only simpler of application than the classical methods, but well adapted for purposes of exposition. The simplicity of the demonstration rests in referring the general question of equivalence of families back to that of the similarity of numerical matrices. The question of similarity is settled by the derivation of a normal form in which a complete system of invariants of a class of matrices is displayed.

§ 1.

Two families of bilinear forms in $2n$ variables,

$$\lambda' A + \lambda'' B = \Sigma(\lambda' a_{ij} + \lambda'' b_{ij}) x_i y_j, \tag{1}$$

$$(i, j = 1, 2, \ldots, n)$$

$$\lambda' A' + \lambda'' B' = \Sigma(\lambda' a'_{ij} + \lambda'' b'_{ij}) x'_i y'_j, \tag{2}$$

are equivalent when and only when linear transformations,

$$x_i = \Sigma_l h_{il} x'_l \qquad |h_{il}| \neq 0,$$

$$y_j = \Sigma_k k_{jk} y'_k \qquad |k_{jk}| \neq 0,$$

exist that will transform (1) into (2). In these families, the parameters λ' and λ'', and the coefficients, are assumed to be real or complex numbers. This transformation may be expressed as follows in the notation of matrices:

$$\mathbf{h}(\lambda'\mathbf{a} + \lambda''\mathbf{b})\mathbf{k} = \lambda'\mathbf{a}' + \lambda''\mathbf{b}', \tag{3}$$

where **h**, **a**, etc. symbolize the matrices whose determinants are $|h_{il}|$, $|a_{ij}|$, etc.

The necessary and sufficient conditions for the validity of (3) are that

$$\mathbf{hak} = \mathbf{a}'; \quad \mathbf{hbk} = \mathbf{b}'. \tag{4}$$

The sufficiency of these conditions follows immediately from the application of the distributive law in operations with matrices; their necessity, from the fact that (3) is valid for all values of λ' and λ''.

We proceed to find a method of determining the conditions under which **h** and **k** exist; that is, simple means of determining whether (1) and (2) are equivalent. Let us first consider the exceptional case when $|\mathbf{a}| \equiv |\mathbf{b}| \equiv 0$; while $|\lambda'\mathbf{a} + \lambda''\mathbf{b}| \equiv |\mathbf{c}| \not\equiv 0$. For a certain pair of values of λ' and λ'' for which the determinant $|\mathbf{c}|$ does not vanish, let $\lambda'\mathbf{a}' + \lambda''\mathbf{b}' \equiv \mathbf{c}'$. Then if **c** and **c**$'$ are equivalent,

$$\mathbf{hck} = \mathbf{c}'; \quad \text{and} \quad \mathbf{hak} = \mathbf{a}'.$$

Consequently

$$\mathbf{k} = \mathbf{c}^{-1}\mathbf{h}^{-1}\mathbf{c}', \quad \text{and} \quad \mathbf{a}' = \mathbf{hac}^{-1}\mathbf{h}^{-1}\mathbf{c}',$$

or

$$\mathbf{a}'\mathbf{c}'^{-1} = \mathbf{h}(\mathbf{ac}^{-1})\mathbf{h}^{-1}; \tag{5}$$

thus if **c** and **c**$'$ are equivalent, the matrix $\mathbf{a}'\mathbf{c}'^{-1}$ is similar* to the matrix $\mathbf{ac}^{-1}$. The converse is also true. For from (5), letting $\mathbf{k} = \mathbf{c}^{-1}\mathbf{h}^{-1}\mathbf{c}'$, we obtain immediately

$$\mathbf{a}' = \mathbf{hak}; \quad \mathbf{c}' = \mathbf{hck}.$$

This criterion also holds in the unexceptional case where at least one of the determinants $|\mathbf{a}|$ and $|\mathbf{b}|$, say $|\mathbf{b}|$, does not vanish. In this case the condition (5) is replaced by the condition

$$\mathbf{a}'\mathbf{b}'^{-1} = \mathbf{h}(\mathbf{ab}^{-1})\mathbf{h}^{-1}. \tag{6}$$

The case where $|\lambda'\mathbf{a} + \lambda''\mathbf{b}| \equiv 0$, which Kronecker has completely discussed, will not be considered in this paper.

The problem of determining the equivalence of families of bilinear forms (1) and (2) is, then, reduced to the problem of determining the similarity of the matrices $\mathbf{a}'\mathbf{c}'^{-1}$ and $\mathbf{ac}^{-1}$. It should be observed that the determinant of these matrices may be zero.

---

* Frobenius defines the matrices **a** and **b** to be similar (ähnlich) when a matrix **l** exists such that $\mathbf{a} = \mathbf{lbl}^{-1}$. See *Crelle*, 84, p. 21.

§ 2.

We may, if we choose, think of all the matrices that are similar to each other as forming a class of matrices. If we can determine a complete system of invariants for a class of matrices, we shall be in a position to settle the question of the equivalence of bilinear forms that was stated in the last section. As a matter of fact, we shall prove that the following constitute such a system for a class of matrices under the transformations of similarity:

I. The rank of any matrix, $\mathbf{a}$, of the class.

II. The characteristic equation of $\mathbf{a}$.

III. The degrees of the elementary divisors of the matrix $\lambda - \mathbf{a}$.

The proof of the first two parts of this statement is very simple and familiar, but will be given for the sake of completeness.

Invariant I. Given a matrix $\mathbf{a}$ of order n and rank k. Let $\mathbf{b} = \mathbf{hah}^{-1}$. Then since the rank of the product of determinants cannot exceed the lowest rank of the factors, the rank of $\mathbf{b}$ cannot be greater than that of $\mathbf{a}$. Similarly, since $\mathbf{a} = \mathbf{h}^{-1}\mathbf{bh}$, and the rank of $\mathbf{a}$ cannot exceed that of $\mathbf{b}$. Hence the rank is invariant.

Invariant II. By the characteristic equation of a matrix

$$\mathbf{a} = \left\Vert \begin{matrix} a_{11} & a_{12} & a_{13} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & a_{32} & \cdots\cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots\cdots & a_{nn} \end{matrix} \right\Vert$$

is meant the equation of the nth degree in λ that one obtains by developing the determinant in the equation

$$\phi(\lambda) = \left| \begin{matrix} \lambda - a_{11} & a_{12} & a_{13} \cdots\cdots\cdots a_{1n} \\ a_{21} & \lambda - a_{22} & a_{23} \cdots\cdots\cdots a_{2n} \\ a_{31} & a_{32} & \lambda - a_{33} \cdots\cdots\cdots a_{3n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots\cdots \\ a_{n1} & a_{n2} & a_{n3} \cdots\cdots \lambda - a_{nn} \end{matrix} \right| = 0.$$

Suppose this equation has the form

$$\phi(\lambda) = \lambda^n + \alpha_1\lambda^{n-1} + \alpha_2\lambda^{n-2} + \cdots\cdots + \alpha_n = 0.$$

The theorem is well-known that the matrix $\mathbf{a}$ satisfies this equation identically, that is, $\phi(\mathbf{a}) \equiv 0$. From this it follows that the characteristic equation is an invariant of the class of similar matrices. For it is only necessary to show that the equation $\phi(\lambda) = 0$ is satisfied identically when $\mathbf{hah}^{-1}$ is substituted for λ. This fact appears immediately, for

$$\phi(\mathbf{hah}^{-1}) = (\mathbf{hah}^{-1})^n + \alpha_1(\mathbf{hah}^{-1})^{n-1} + \dots\dots + \alpha_n$$
$$= \mathbf{ha}^n\mathbf{h}^{-1} + \alpha_1\mathbf{ha}^{n-1}\mathbf{h}^{-1} + \dots\dots + \alpha_n = \mathbf{h}\phi(\mathbf{a})\mathbf{h}^{-1} \equiv 0.^*$$

Invariant III. We now proceed to define the elementary divisors of the λ-matrix $\lambda - \mathbf{a}$. As we have seen, the determinant of this matrix is a rational integral equation in λ, and is consequently factorable into n linear factors, $(\lambda - \lambda_1)^{l_1}(\lambda - \lambda_2)^{l_2}(\lambda - \lambda_3)^{l_3} \dots\dots (\lambda - \lambda_k)^{l_k}$, where $l_1 + l_2 + l_3 + \dots\dots + l_k = n$. Suppose that $(\lambda - \lambda_1)^{l'_1}$ is the highest power of $\lambda - \lambda_1$ that is contained in all the first minors of the determinant $|\lambda - \mathbf{a}|$, that l''_1 is the highest exponent of this same factor that is found in all the second minors, and that, continuing in the same way, $l_1^{\tau_1}$ is the highest exponent of $\lambda - \lambda_1$ that is found in all the τth minors of the determinant. Similarly, we may obtain sets of exponents bearing the same relation to the other factors of the original determinant, such as

$$l'_2,\ l''_2,\ l'''_2, \dots\dots,\ l_2^{\tau_2};\ \dots\dots;\ l'_k,\ l''_k,\ l'_k,\ \dots\dots,\ l_k^{\tau_k}.$$

Now the first differences of these exponents,

$$l_1 - l'_1 = e_1,\ l'_1 - l''_1 = e_2,\ l''_1 - l'''_1 = e_3,\ \dots\dots,\ l_1^{\tau-1} - l_1^{\tau} = e_\tau,$$

are called the exponents of the elementary divisors of the matrix. The expressions $(\lambda - \lambda_1)^{e_1}$, $(\lambda - \lambda_1)^{e_2}$, $\dots\dots$, $(\lambda - \lambda_1)^{e_\tau}$ are called the elementary divisors of the matrix $\lambda - \mathbf{a}$, corresponding to the factor $\lambda - \lambda_1$. We can obtain the complete set of elementary divisors of the matrix by considering the other factors of the characteristic equation.

We shall, however, be interested chiefly in the exponents e_1, e_2, e_3, $\dots\dots$ It appears that if the roots of the characteristic equation and these e's, together with their distribution over the various factors, are known, we then know all the elementary divisors of the matrix. If, then, we prove that the elementary divisors are invariants of a class of λ-matrices, the only actual addition to our two invariants I and II are these e's and their distribution on the various factors of the characteristic equation.

The proof that the elementary divisors of a matrix are invariants of the matrix follows from the theorem that the rth minors of the product of two

---

[*] Of course this same demonstration would show that any equation satisfied by $\mathbf{a}$ is an invariant of the class of matrices.

determinants are linear in each of the rth minors of the factors. Hence the elementary divisors of the matrix $\lambda - \mathbf{a}$, and of any matrix equivalent to it, as $\mathbf{s}(\lambda - \mathbf{a})\mathbf{t} = \lambda - \mathbf{b}$, are the same.

Now if $\mathbf{a}$ and $\mathbf{b}$ are similar, $\lambda - \mathbf{a}$ and $\lambda - \mathbf{b}$ are equivalent, and conversely. For suppose that $\mathbf{a}$ and $\mathbf{b}$ are similar. Then the matrices $\lambda - \mathbf{a}$ and $\lambda - \mathbf{b}$ are equivalent. For if $\mathbf{b} = \mathbf{hah}^{-1}$, then evidently $\lambda - \mathbf{b} = \mathbf{h}(\lambda - \mathbf{a})\mathbf{h}^{-1}$, since $\mathbf{h}\lambda\mathbf{h}^{-1} = \lambda$. Conversely, if $\lambda - \mathbf{a}$ and $\lambda - \mathbf{b}$ are equivalent, then $\mathbf{a}$ and $\mathbf{b}$ are similar. For, suppose that $\mathbf{s}(\lambda - \mathbf{a})\mathbf{t} = \lambda - \mathbf{b}$. Since the distributive law holds in the operations with matrices, we have $\mathbf{s}\lambda\mathbf{t} - \mathbf{sat} = \lambda - \mathbf{b}$, or since this equation is identically true, $\mathbf{s}\lambda\mathbf{t} = \lambda$, that is, $\mathbf{st} = 1$, or $\mathbf{t} = \mathbf{s}^{-1}$. Hence $\mathbf{b} = \mathbf{s}^{-1}\mathbf{as}$.

Consequently in determining whether the matrices mentioned at the end of § 1 are similar or not, it is only necessary to discuss the equivalence of the corresponding λ-matrices.

§ 3.

This completes the demonstration that I, II and III are invariants of a class of matrices. It is a more serious matter to show that they form a complete system of invariants. To do this, we shall show how to reduce the matrix $\mathbf{a}$ into a similar normal form which is completely characterized by the system of invariants I, II and III.

Let the roots of $|\lambda - \mathbf{a}| = 0$ be $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$, of which suppose that the first r are distinct. Let the solutions of the r sets of n linear equations in n variables

$$\sum_k a_{lk}x_k = \lambda_h x_l \quad (l = 1, 2, \ldots, n;\ h = 1, 2, \ldots, r)$$

be $\xi_1^{(h)}\ \xi_2^{(h)}\ \xi_3^{(h)} \ldots \xi_n^{(h)}$. Build, now, a matrix of the nth order of which the first r columns are these ξ's, and of which the other elements are all zero excepting those in the principal diagonal, which is made up of 1's in the places not filled with ξ's. Call this matrix $\mathbf{t}$. Build also $\mathbf{t}^{-1}$, and carry out the multiplication $\mathbf{t}^{-1}\mathbf{at}$. The result of this process by which we pass from $\mathbf{a}$ to a similar matrix changes the form of the matrix materially. In fact, in the new form we have all the elements of the first r columns zero except the elements of the principal diagonal, which are $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_r$. In the transformed matrix consider the $(n - r)$-rowed determinant that is common to the last $n - r$ rows and columns. Proceed with this exactly as was done with the original matrix, only in the matrix that corresponds to $\mathbf{t}$ let the elements in the first r columns all be zero except those in the principal diagonal that fall in those columns, which we place equal to 1.

Continue this process till there are no non-zero elements below the principal diagonal, and the principal diagonal consists of the roots of the characteristic equation.

Numerical Example. Let us reduce to normal form the matrix

$$\mathbf{a} = \left\| \begin{array}{rrrr} 2 & -2 & 1 & 1 \\ 6 & -3 & 0 & 4 \\ 16 & -3 & -4 & 8 \\ 23 & 4 & -13 & 10 \end{array} \right\|$$

of which the characteristic equation is

$$\lambda^4 - 5\lambda^3 + 9\lambda^2 - 7\lambda + 2 = 0.$$

By taking the derivative of this equation it appears that 1 is a triple root, and that 2 is the other root.

We will first simplify the matrix by bringing 2 to the leading place, and 1 to the next place in the principal diagonal, making all the other elements of the first two columns zero. To this end we get as solutions of the equations

$$\begin{aligned} 2x - 2y + z + w &= 2x, \\ 6x - 3y \phantom{{}-4z} + 4w &= 2y, \\ 16x - 3y - 4z + 8w &= 2z, \\ 23x + 4y - 13z + 10w &= 2w \end{aligned}$$

the values $x = 1$, $y = 2$, $z = 3$, $w = 1$. We must also solve the equations that we obtain from these by replacing the right-hand members, viz., $2x$, $2y$, $2z$, $2w$, by x, y, z, w. The solution of these equations is $x = 0$, $y = -1$, $z = -1$, $w = -1$. Now compute the product $\mathbf{t}^{-1}\mathbf{a}\mathbf{t}$, where

$$\mathbf{t} = \left\| \begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 3 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array} \right\|, \quad \mathbf{t}^{-1} = \left\| \begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array} \right\|,$$

and we obtain

$$\mathbf{a} = \left\| \begin{array}{rrrr} 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & -5 & 3 \\ 0 & 0 & -12 & 7 \end{array} \right\|.$$

Now solving the equations

$$\begin{aligned} -\ \ 5x + 3y &= x, \\ -12x + 7y &= y, \end{aligned}$$

we find that $x = 1$, $y = 2$. Now compute $\mathbf{t}_1^{-1}\mathbf{a}\mathbf{t}_1$, where

$$\mathbf{t}_1 = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{Vmatrix},$$

and we obtain

$$\mathbf{a} = \begin{Vmatrix} 2 & 0 & 1 & 1 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

From the results that we shall obtain in the next paragraphs it appears that we may replace the last two elements in the first row, and the last element in the second row, by 0 without further transformation. The -2 and the 3 may be replaced by 1, and we have the result

$$\mathbf{a} = \begin{Vmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

Now if a non-zero element, say a_{ik}, occurs in the intersection of the ith row, and the kth column, and if the roots λ_i and λ_k are distinct, this element may be replaced by zero without otherwise disturbing the matrix. For build the matrices $\mathbf{h}$ and $\mathbf{h}^{-1}$ each of which have all the elements in their principal diagonal equal to 1, and which have all the other elements equal to zero except that in the ith row and the kth column, where the values h and $-h$ respectively occur, where $h = \frac{a_{ik}}{\lambda_i - \lambda_k}$. The transformation $\mathbf{h}^{-1}\mathbf{a}\mathbf{h}$ does not change the elements of any row above or below the ith, and only the element a_{ik} in that row, which it replaces by 0.

Consequently, since we can always find a transformation that will replace such elements a_{ik}, $(\lambda_i \neq \lambda_k)$ by zero without disturbing the rest of the matrix, there is no necessity of going through the transformation at all, but the elements may be replaced by zero as soon as the previous transformations are completed.

We may now, by an interchange of rows and columns, which in the present form of matrix is equivalent to passing to a similar matrix, bring into juxtaposition along the principal diagonal all the roots that are equal to each other. We have then a matrix of the form

$$\mathbf{a}=\left\Vert\begin{array}{cccc} \mathbf{a}_1 & 0 & 0 \;\ldots\ldots & 0 \\ 0 & \mathbf{a}_2 & 0 \;\ldots\ldots & 0 \\ 0 & 0 & \mathbf{a}_3 \ldots\ldots & 0 \\ \ldots & \ldots & \ldots\ldots\ldots & \ldots \\ 0 & 0 & 0 \;\ldots\ldots & \mathbf{a}_k \end{array}\right\Vert, \tag{7}$$

where $\mathbf{a}_j$ represents a submatrix all of whose roots are equal, and in which all the elements below the principal diagonal are zero.

§ 4.

We may then for the further reduction of the matrix confine ourselves to the consideration of a matrix all of whose roots are equal, and which is in the form

$$\mathbf{a}_1=\left\Vert\begin{array}{ccccc} \lambda & a_{12} & a_{13} & \ldots\ldots & a_{1m} \\ 0 & \lambda & a_{23} & \ldots\ldots & a_{2m} \\ 0 & 0 & \lambda & \ldots\ldots & a_{3m} \\ \ldots & \ldots & \ldots & \ldots\ldots & \ldots \\ 0 & 0 & 0 & \ldots\ldots & \lambda \end{array}\right\Vert,$$

where λ is taken to represent any one of the distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_k$. We now seek a method of further reducing matrices of this type. To effect this reduction we may use transformations of three types.

First, we may effect an interchange of the subscripts i and j; that is, we may carry out the substitution (ij) on the subscripts by transforming $\mathbf{a}_1$ by the matrix $\mathbf{h}$ in which $h_{ij}=h_{ji}=1=h_{kk}$, $(k \neq i, j)$, and all the other elements are zero. Since the transformation is known always to exist, of course the indices may be interchanged without actually going through the transformation.

Second, if $a_{ij} \neq 0$, and at the same time $a_{j-1j} \neq 0$, we may transform $\mathbf{a}_1$ by $\mathbf{h}$, where $h_{kk}=1$ and all the other elements are zero excepting h_{ij-1}, which has the value a_{ij}/a_{j-1j}. This transformation has the effect of replacing a_{ij} by zero, but not affecting the remainder of the matrix excepting certain elements in the first $i-1$ rows.

Third, if $a_{ij} \neq 0$, and at the same time $a_{j-1j}=0$, but $a_{ii+1} \neq 0$, we may transform the matrix $\mathbf{a}_1$ by a transformation of the first type so that a_{ij} falls in the place that was occupied by a_{ii+2}, and is, as a consequence, directly over a zero element. Now transform the matrix by a transformation $\mathbf{h}$ such that $h_{kk} = 1$ and all the other elements are zero excepting $h_{i+2\,i+1}$, which has the value $-a_{ii+1}/a_{ij}$. The effect of this transformation is to replace a_{ii+1} by zero, and to leave all the other elements in the matrix unchanged with the exception of certain elements that occur in the first $i-1$ rows.

By the use of these transformations we are able to start with the last row, and gradually replace all the elements by zero, excepting certain elements that we can always put in the place of the elements a_{ii+1}, and which constitute what we will call the over-principal diagonal. By a simple transformation these elements may all be replaced by unity, so that our matrix $\mathbf{a}_1$ assumes a form in which every element is zero excepting the elements of the principal diagonal, which are all equal to λ, and the elements of the over-principal diagonal, which are all either 1 or 0. This is the normal form for which we were seeking. It appears that any matrix can be brought to this form by properly chosen transformations, and consequently that every matrix is similar to some matrix which is in normal form. It is observed that the matrices that we have considered need not necessarily have the rank n, but the characteristic equation may have zero roots without affecting the method in the least.

The transformation employed in obtaining the form (7) assumed that the characteristic equation had been solved. If this equation is irreducible in the domain of rational numbers, this process is an irrational one. Since the objection to the irrational process is partly a practical one, it is worth while to observe how serious these irrational operations are, practically. In the first place, if there are no multiple roots in the characteristic equation, the fact is determined rationally, and in such a case no reduction is ever necessary, as each of the subdeterminants $\mathbf{a}_i$ in (7) reduces to a single element, and the equivalence hinges on the identity of the characteristic equations themselves, which is surely determined by a rational process. For the case $n = 3$ there can be no irrational multiple root. In the case $n = 4$, the only possibility of having an irrational multiple root is when the characteristic equation is a perfect square; that is, when there are two pairs of conjugate quadratic surd or complex roots. No irrationalities of the third order will be encountered in the reduction of any matrix of lower order than the ninth, and then only when the characteristic equation is a perfect cube; and a matrix of order twenty-five would be the first that would involve non-algebraic irrationalities.

The submatrices in the previous numerical example that correspond to equal roots of the characteristic equation are so simple that it is worth while to give another numerical example of a rather more complicated character. In considering the following matrix of order six, with six equal roots, it is to be noted that the value of λ is assumed to be 1.

$$\mathbf{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & -3 & 5 & -7 \\ 0 & 1 & 1 & -1 & 2 & -5 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Transforming by $\mathbf{h}$ where all the elements of $\mathbf{h}$ are zero except those in the principal diagonal, and the element h_{65}, all of which have the value 1, we obtain

$$\mathbf{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & -3 & -2 & -7 \\ 0 & 1 & 1 & -1 & -3 & -5 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

This is a transformation of the third type.

Interchanging, now, the elements that should be moved when the subscripts 5 and 6 are interchanged, we have

$$\mathbf{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & -3 & -7 & -2 \\ 0 & 1 & 1 & -1 & -5 & -3 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Transforming this matrix by $\mathbf{h}$ where all the elements of $\mathbf{h}$ are zero excepting those in the principal diagonal, and the element h_{34}, all of which have the value 1, we obtain

$$\mathbf{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & -2 & -7 & -2 \\ 0 & 1 & 1 & 0 & -5 & -3 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

This and the following transformation are of the second type.

Transforming this matrix by $\mathbf{h}$, where all the elements of $\mathbf{h}$ are zero excepting those in the principal diagonal, which have the value 1, and h_{24} which has the value 5, we obtain

$$\mathfrak{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & 3 & -7 & -2 \\ 0 & 1 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Interchanging the elements corresponding to an interchange of the subscripts 4 and 6, followed by the interchange of 5 and 6, we obtain

$$\mathfrak{a}_1 = \begin{Vmatrix} 1 & 1 & 1 & -2 & 3 & -7 \\ 0 & 1 & 1 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Transforming this matrix by $\mathbf{h}$, where the principal diagonal consists of 1's and $h_{43} = \frac{1}{3}$, we get

$$\mathfrak{a}_1 = \begin{Vmatrix} 1 & 1 & \frac{1}{3} & -2 & -3 & -7 \\ 0 & 1 & 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

This is a transformation of the third type.

From the nature of the transformations employed it is clear that we may get rid of elements without disturbing the rest of the matrix, so that the normal form of the matrix is finally

$$\mathfrak{a}_1 = \begin{Vmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

It is easy to see that the necessary and sufficient condition that two matrices are similar is that they are reducible to the same normal form. Furthermore, it appears that if we know the rank of a matrix and the roots of its characteristic equation, that is if we have the first two of the invariants mentioned in § 2, we can fix exactly every feature of the normal form of that matrix excepting the arrangement of the 1's along the over-principal diagonal. If, then, we can show that the way these group themselves is determined by the values of the exponents of the elementary divisors, we shall be able to pronounce the set of three invariants a complete system for the class of matrices.

Let the matrix $\mathbf{a}_1$ have the following form:

$$\begin{array}{l}
\left.\begin{array}{l}\lambda 1\\ \quad\lambda 1\\ \qquad\cdot\;\cdot\end{array}\right\rangle e_1\\
\qquad\quad\lambda 0\\
\qquad\qquad\left.\begin{array}{l}\lambda 1\\ \quad\lambda 1\\ \qquad\cdot\;\cdot\end{array}\right\rangle e_2\\
\qquad\qquad\qquad\quad\lambda 0\\
\qquad\qquad\qquad\qquad\left.\begin{array}{l}\lambda 1\\ \quad\lambda 1\\ \qquad\cdot\;\cdot\end{array}\right\rangle e_k\\
\qquad\qquad\qquad\qquad\qquad\quad\lambda
\end{array}$$

where zeros fill the places left blank, and where the e's represent the number of consecutive 1's in the various groups of 1's in the over-principal diagonal. Suppose that the order of subscripts is so chosen that $e_1 > e_2 > \dots > e_k$. Then we can show that these e's are precisely the exponents of the elementary divisors that constitute one of the three invariants mentioned.

Since the matrix $\lambda - \mathbf{a}_1$ only differs from the matrix $\mathbf{a}_1$ in having the elements λ replaced by the elements $\lambda - \lambda_1$ all along the principal diagonal, we may observe the relation between the elementary divisors and the way the 1's are arranged as well from the matrix $\mathbf{a}_1$ that we have just written as from the matrix $\lambda - \mathbf{a}_1$. We wish to observe the highest common factor of the subdeterminants of the various orders of $\mathbf{a}_1$. Consider the highest power of $\lambda - \lambda_1$ that occurs in all the non-vanishing first minors of $\lambda - \mathbf{a}_1$. Evidently the first minors with respect to the elements above the principal diagonal vanish, since all the elements below the principal diagonal of such a minor, and at least one element in the principal diagonal, are zero. The minor with respect to an element in the principal diagonal has the value $(\lambda - \lambda_1)^{m-1}$, where m is the order of $\mathfrak{a}_1$. If in forming a first minor with respect to an element below the principal diagonal, one of the zeros that separate the sets of consecutive 1's is inside the right angle formed by the row and column that are deleted to form the first minor in question, then that zero is introduced into the principal diagonal of the first minor, which, therefore, vanishes. Consequently, the only elements of $\mathfrak{a}_1$ whose minors do not vanish are those whose rows and columns contain only consecutive 1's from the over-principal diagonal of $\mathfrak{a}_1$ inside the right angle formed by those rows and columns. Consequently the minimum degree of any non-vanishing first minor in $\lambda - \lambda_1$ is $m - e_1$. Similarly the minimum degree of $\lambda - \lambda_1$ in the 2nd, 3rd, ... minors is $m-(e_1+e_2)$, $m-(e_1+e_2+e_3)$, ... Evidently the first difference of these numbers are $e_1, e_2, e_3, \ldots$; that is, the e's are the exponents of the elementary divisors with respect to the root λ_1. In a precisely similar manner we could find corresponding numbers for the other roots of the characteristic equation of the original matrix $\mathfrak{a}$. The e's of our normal form are, therefore, the exponents of the elementary divisors of the matrix.

We have, as the result of the paper up to the present, the

THEOREM. The necessary and sufficient condition that the two non-singular families of bilinear forms

$$\lambda' A + \lambda'' B, \tag{1}$$

$$\lambda' A' + \lambda'' B' \tag{2}$$

are equivalent, is that the matrices corresponding to the forms

$$\lambda - AB^{-1},$$

$$\lambda - A'B'^{-1}$$

have, as their complete system of invariants under the similarity transformation, (I) their order, (II) the roots of their characteristic equations, (III) the exponents of their elementary divisors.

These three invariants may be included in one invariant; namely, the elementary divisors of the matrices of these forms.

§ 5.

We proceed to show that the necessary and sufficient condition that the two non-singular families $\lambda' A + \lambda'' B$ and $\lambda' A' + \lambda'' B'$ are equivalent is that their own elementary divisors are identical. The necessity of this condition is shown in the same way as the invariance of the elementary divisors of a matrix was shown at the end of § 2, and will not be repeated here.*

It also appears immediately that if the matrices $\mathfrak{a}\mathfrak{b}^{-1}$ and $\mathfrak{a}'\mathfrak{b}'^{-1}$ have identical elementary divisors, then (1) and (2) also have identical elementary divisors. For then $\mathfrak{a}\mathfrak{b}^{-1}$ and $\mathfrak{a}'\mathfrak{b}'^{-1}$ are similar, hence (1) and (2) are equivalent, and have the same elementary divisors. It only remains to show that if (1) and (2) have equal elementary divisors, then $\mathfrak{a}\mathfrak{b}^{-1}$ and $\mathfrak{a}'\mathfrak{b}'^{-1}$ also have equal elementary divisors.

Let $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{n}$, and consider the family $\lambda_1\mathfrak{n} + \lambda_2$. We assume that $\mathfrak{n}$ is in normal form. This family is equivalent to (1) by the criterion derived in § 1.

Now if the matrix of this family $\lambda_1\mathfrak{n} + \lambda_2$ is written out, it is observed that the principal diagonal consists of elements of the form $\lambda_1\lambda_i + \lambda_2$, where the λ_i are roots of the characteristic equation of the matrix $\mathfrak{a}\mathfrak{b}^{-1}$, and the over-principal diagonal consists of unit elements or zeros, in the same order as they occur in the over-principal diagonal of the matrix $\mathfrak{n}$. That is to say, we have here a form of the matrix of the family of (1) that displays its elementary divisors in just the same way as the elementary divisors of $\mathfrak{a}\mathfrak{b}^{-1}$ are displayed in $\mathfrak{n}$. It is, in fact, a normal form for (1), and it appears that the exponents of the elementary divisors for $\mathfrak{n}$ are the same as those for (1). Consequently, when we find the normal form for $\mathfrak{a}\mathfrak{b}^{-1}$, we are simultaneously finding the normal form (1), and if (1) is known to have certain elementary divisors, we can write down without further discussion the elementary divisors of $\mathfrak{a}\mathfrak{b}^{-1}$, for their exponents are identical. Consequently, if (1) and (2) have the same elementary divisors, $\mathfrak{a}\mathfrak{b}^{-1}$ and $\mathfrak{a}'\mathfrak{b}'^{-1}$ also have the same elementary divisors, and (1) and (2) are therefore equivalent. This is Weierstrass's great theorem.

Practically to determine whether two families of bilinear forms of the types (1) and (2) are equivalent, we must build the matrices $\mathfrak{a}\mathfrak{b}^{-1}$ and $\mathfrak{a}'\mathfrak{b}'^{-1}$, and reduce both of these matrices to their normal form. If they have identical (I) rank, (II) characteristic equations, (III) exponents of their elementary divisors, then (1) and (2) are equivalent. It is to be observed that if invariants (I) or (II) of the two matrices differ, the fact is always noticed before any reduction is made, and the fact of inequivalence is shown without the labor of reduction.

YALE UNIVERSITY, *December* 11, 1908.

---

*For a detailed proof of this fact see Bôcher, Introduction to Higher Algebra.